Mathematical Metaphor

DAVID PIMM

> In order to make a possible distinction between metaphorical and literal language, the definition of metaphor must be pushed one step further. It is not only the extension of a word to a novel referent on the grounds of similarity that constitutes a metaphor: it is the deliberate extension of this word *across* previously established category boundaries [Winner, 1979, p. 472]

Metaphor involves the seeing (and therefore the understanding) of one thing in terms of another; it is a conceptual rather than solely a lingustic phenomenon. One of the available metaphors for mathematics itself is that of seeing mathematics as a language. One particular instance of this metaphor at work arises when we talk of "the mathematical literature". If we take this latter expression as a normal adjectival one, then it appears a commonplace deduction that mathematical literature *is* literature. One consequence of this apparently trivial claim is the possibility of applying some of the insights and techniques of literary criticism to mathematical text

Rhetoric

Lakatos' book *Proofs and refutations* [1976] provides one instance of what might be called mathematical literary criticism In it, he explored some of the conventions of mathematical activity and developed a set of concepts (such as monster barring, local and global counterexamples, or lemma incorporation) for talking about mathematics — meta-mathematical concepts in some sense. In passing, I would like to reclaim the term meta-mathematics for mathematics education and to assert the power of metaphors for mathematics such as "mathematics as language" in terms of their being able to cast unexpected light on mathematics. They thereby enable questions to be posed that may not previously have been seen as relevant or even coherent. [See Pimm, 1987 for further discussion of this point.]

Applying stylistic notions to mathematical text can produce some insight into the possible intentions of the authors, as can the examination (so beloved of semioticians) of the relations among reader, writer and the text itself. To take one particular example, consider the presence (or otherwise) of the author in mathematical text and the concomitant acknowledgement of and attitude to the reader. The mathematician David Fowler, in the introduction to his book *Introducing real analysis,* comments, "I, the author, address you, the reader, in a way that may be considered unseemly by my colleagues". Elsewhere [Pimm, 1984] I have written about the use of the pronoun "we" in mathematical discourse and mused about to whom or which community this word is referring. For instance, consider the following examples from a mathematics textbook by Fraleigh [1967, my emphasis].

> As mathematicians, let *us* attempt . . .
>
> *We* have seen that any two groups of order three are isomorphic. *We* express this by saying that there is only one group of order three up to isomorphism.
>
> In the past, some of the *author's* students have had a hard time understanding and using the concept of isomorphism. *We* introduced it several sections before *we* made it more precise in the hope that *you* would really comprehend the importance and meaning of the concept. Regarding its use, *we* now give an outline showing how the mathematician would proceed from the definition . . .

Fauvel [1988, this issue] has distinguished between what he calls the *Euclidean* and *Cartesian* forms of mathematical rhetoric.

> Euclid's attitude is perfectly straightforward: there is no sign that he notices the existence of readers at all. . . The reader is never addressed.

He claims that the closest Euclid comes is the imperative mood of apparent requests ("let such-and-such be done"), but they are validated by the postulates seen as abstract permissions, and certainly not as instructions or requests to the reader

Fauvel contrasts this with Cartesian rhetoric as follows:

> Instead of a clear, linear mathematical account in a Euclidean form, Descartes put out an extraordinary blend of hints, procedures, assertions, truths and falsehoods — a carefully controlled stream of consciousness.

Of the style of the *Discourse on Method* as a whole, he earlier writes:

> The *Discourse* turns out to be a finely constructed story about the past persona (called "I") of a narrator (also called "I"), structured so as to bring out an

For the Learning of Mathematics **8**, 1 (February 1988)
FLM Publishing Association Montreal, Quebec, Canada

imaginary intellectual journey — a fictional narrative cast in the form of an autobiography

He ends by asserting that by means of singling out different styles, we suddenly notice that there are different rhetorical forms in mathematical exposition, that the choice is under the author's control and that a greater awareness of the range of such forms may contribute to an understanding of how we communicate in mathematics. In particular, by drawing attention to alternative rhetorical possibilities, the hegemony of the Euclidean style as well as the reasons for taking that form may be open to scrutiny.

Just as the claim that science is value-free needs to be exploded, so the claim that mathematics is rhetoric-free is part of an ideology of mathematics — one which can be related to the realism debate in literary styles. Thus, just as so-called "realistic" writing is as much a style with its own conventions worthy of study in relation to its effects as any other, so is the Euclidean rhetorical style. Leron [1983, 1985] has written two articles that might also be identified as mathematical literary criticism in which he has offered an alternative to the "linear" style of proof exposition which forms part of the Euclidean rhetorical style.

Metaphor and metonymy

The traditional study of rhetoric offers two concepts, metaphor and metonymy, as the main figures or tropes with which to explore literary form and meaning. In this section, by means of these notions, I explore some of the ways in which mathematical meaning is coded into particular expressions. The linguist Michael Halliday has coined the term *register* which he uses to describe:

> a set of meanings that is appropriate to a particular function of language, together with the words and structures which express these meanings. . . We can refer to a "mathematics register", in the sense of the meanings that belong to the language of mathematics (the mathematical use of natural language, that is: not mathematics itself), and that a language *must* express if it is used for mathematical purposes. . . In order to express new meanings, it may be necessary to invent new words; but there are many different ways in which a language can add new meanings, and inventing new words is only one them. [Halliday, 1975, p. 65]

This article forms part of a larger enterprise to explore the nature of the mathematics register in English.

Lakoff and Johnson [1980, p. 125] claim that "metaphors are systematic devices for further defining a concept and for changing its range of applicability." I hope to illustrate the force of this quotation in the course of this section with both mathematical and non-mathematical examples, as well as indicate one particular linguistic structure of English by which metaphors are created.

There is a fundamental metaphoric structure in English. It has a very simple linguistic structure, namely

(Article) (Adjective) Noun (*)

so let me give some examples:

a genetic fingerprint
electronic mail
a chemical messenger
a molecular sieve
a chemical equation
a technology
a spherical triangle
the complex plane
matrix multiplication
differential geometry
an imaginary sphere
a geometry

When first coined, the adjective forms a novel collocation with the particular noun, and as we were taught in school, an adjective modifies a noun. The modification in these cases, however, is considerable. The adjective points to a new context of application, and sometimes considerable intellectual effort is required in order to create a meaning for the whole expression.

Consider first the expression "a genetic fingerprint". The construction suggests the adjective is modifying the noun in the customary fashion and therefore it is true (if tautological), based on a knowledge of English, that a genetic fingerprint *is* a fingerprint. Russell [1970, p. 46], among many others, wrote about the relation between *a priori* and "analytic" truths.

> Before the time of Kant, it was generally held that whatever knowledge was *a priori* must be "analytic". What this word means will be best illustrated by examples. If I say, "A bald man is a man", "A plane figure is a figure", "A bad poet is a poet", I make a purely analytic judgement: the subject spoken about is given as having at least two properties, of which one is singled out to be asserted of it. Such propositions as the above are trivial, and would never be enunciated in real life except by an orator preparing the way for a piece of sophistry.

While bearing in mind Russell's warning (signalled by the final sentence) about my intentions, I should nonetheless like to propose this as a criterion for establishing the degree of metaphoricity for an individual in such expressions, namely the difficulty with which such ostensibly trivial, "analytic" claims can be construed.

Initially, at the literal level of meaning, a genetic fingerprint (assuming the process to which the expression refers is familiar) is *not* a fingerprint. Its use suggests a search for an extended or altered sense for "fingerprint", one which restores the earlier assertion to a triviality. (The fact that such neologistic expressions sometimes have single quotation marks around the noun, but never around the adjective, stresses that it is the semantics of the noun that is under perturbation.)

My reconstruction is as follows. A *fingerprint* refers to a mark made by the finger, but one of the strong connotative meanings of fingerprints is the use to which it is put, namely as a unique identifier of a human individual. (Hence the Paul Simon song about "the myth of fingerprints".) An extended sense of *fingerprint* might be to shift the centre of gravity of the concept to this aspect and broaden it so that the term "fingerprinting" now refers to any process which allows such a distinction to be made. (There might be some scope restrictions retained to preserve the bodily aspect of the notion, so that in a world where everyone had a unique number assigned to them, this would not be referred to as a fingerprint.)

With this altered sense of fingerprint, a genetic fingerprint becomes one type of fingerprint, but now another adjective is needed to distinguish which sort of redundancy grounds (which are sometimes less apparent in foreign expressions, e.g. the *hoi polloi*), "a finger fingerprint" would serve, or perhaps "a digital fingerprint", were that adjective not already appropriated for computer use. If, as may well happen, "digital" fingerprinting were to become an outmoded technology, the main or default meaning of the word "fingerprint" might become that of "genetic fingerprint", and so a complete change in the meaning of the word "fingerprint" might be achieved.

This possibility is being realised with my second example, "electronic mail". In England, mail (or post) still comes hand-delivered through a letter-box. For certain sections of the Californian population, *mail* comes through the computer, and *physical mail* is what the postal service brings. It is even possible that the expression "physical mail" will come to be seen as metaphorical; a complete reversal of field and ground. As many of our communicative (and other) functions are taken over by the computer, there may well be a widespread shift in the meanings of words such as "mail", and an increasing use of the adjective "physical".

The other non-mathematical examples on my list are of the same sort (see Note), and the reader might care to spend a little time working through them and deciding on an altered sense of the noun which allows the usage to be construed as a simple adjectival one. (I won't say any more about the examples like "a geometry", but there is a very important historical shift that came about with the move from talking about "geometry" to talking about "a geometry".)

Where does metonymy come into this account? One common feature across these examples is the shift of focus from the object itself as referent to the function achieved by that object (for instance, mail to communication, messenger to informative, sieve to sieving) and hence to allow for a different object or process to achieve that same end. It is this substitution, of ends for means, that is metonymic, and so this complementary literary trope provides the basis for concept extension and therefore sits inside the metaphoric account.

Metonymy, unlike metaphor, is a linguistic process that is more to do with naming. The Concise Oxford Dictionary [Sykes, 1982] defines it as "The use of the name of one thing for that of another with which it is associated. Metonymy can be described as a process of substitution, usually based on some sort of contiguity in space or time". Thus when a waitress says "the orange juice wants his bill", part of the person's order is being used as a way of referring to the person who ordered it (and in a way which offers great saliency to a restaurant, where the primary consideration is what you had, rather than whether or not you are bald, for instance). Such a designation may enable the bill to be found more quickly, perhaps. Despite the claim that "you are what you eat", there is no sense that the man is being seen as the orange juice — it is a *substitution* of names, X *for* Y, rather than a metaphor, seeing X *as* Y. The basis for metonymic substitution is usually "more obvious than is the case with metaphoric ones, since it usually involves physical or causal associations" [Lakoff and Johnson, 1980, p. 89].

Hence are some further everyday examples of metonymy based mainly on the account in Lakoff and Johnson.

> Those lands belong to the crown.
> The Times hasn't arrived at the press conference.
> Acrylic has taken over the art world.

Part for whole (synecdoche)

> We need some new blood
> All hands on deck

Object used for user

> The buses are on strike

Place for event

> I'm afraid Nicaragua will become another Vietnam

Producer for product

> He bought another Timex
> The Picasso sold for 23 million pounds

I now turn to some mathematical examples of the metaphoric structure (*) and see if a similar analysis holds. My first one is the expression "a spherical triangle". Just as with "genetic fingerprint", it is initially hard to see that a spherical triangle *is* a triangle, certainly in the same sense that an isosceles one is. Just as genetic fingerprinting seems to have nothing to do with fingers, so spherical triangles, once the referent is clear, seems to have little to do with straight-sided triangles. In order to make sense of the expression "spherical triangle", an extended meaning for the concept of triangle is required whereby geodesic replaces straight line (a particular case of the concept of geodesic in the case of the plane), and hence can be sensibly applied to the surface of a sphere. There are no straight lines on a sphere, but there are paths of shortest distance.

The metonym lying at the heart of this metaphoric expression is not one of ends for means — at least not straightforwardly so. It comes through the singling out of the "shortest length" aspect of straight lines and using *that* as the basis for broadening the concept reference. Just as with the "orange juice" example, one particular aspect of the original concept setting has been singled out as particularly salient and used as the basis for the extended sense.

To pick up the earlier comment made by Halliday about the possibility of coining new words, what is the force of calling spherical triangles *triangles?* Such geometric configurations of curves ("lines") on spheres may not have had a descriptive term previously. Classifying them as triangles results in stressing the function of this configuration (that is, three segments of great circles meeting pairwise in three points) in the study of the geometry of the sphere, likening it to the role played by the concept of triangle in the plane. Immediately whole theories, comprising definitions, concepts and theorems line up for examination, "translation" and interpretation. In the process, the adjective "spherical" becomes as appropriate or acceptable as isosceles as a classifier for a type of triangle.

For my second example I have chosen the expression "the complex plane". To most people, the statement "the complex plane is a plane" would be a commonplace one, similar to Russell's examples not worthy of remark. Yet this metaphoric naming and the above identification which underlies it has certain mathematical effects, as can be seen by suggesting that "the complex line" is an alternative expression for the same set which stresses certain different features and ignores others. Referring to C as "the complex plane" highlights the two-dimensional (ordered pair) representation of complex numbers $a + ib$ and encourages an approach to complex analysis by means of two-dimensional real analysis via "real and imaginary parts". The plane metaphor also encourages the seeing "a complex number as a vector" [Nolder, 185], which offers an effective image for complex addition and subtraction, but one less so for multiplication and division.

One aspect of seeing the complex plane as a plane is that it suggests it is two-dimensional, yet in the vector-space sense the complex plane is one-dimensional (over itself). It is possible to get from any non-zero complex number to any complex number by complex multiplication (an expression which is itself another metaphor of this class) — a property shared by the real line; hence "the complex line". This is one argument why the complex plane is not a plane, and so this expression needs to retain some of its metaphorical status otherwise confusion may occur.

"Matrix multiplication" provides a third instance, but I shall not go through the details here. The metonymy arises from the decision about which properties associated with whole-number multiplication can be seen as a form of multiplication. Again, this is not an ends-for-means substitution, rather a move from particular operation and setting to characterisation by property. One question for the novice is why these and not other properties were chosen from the available set of perceptions and associations. (I can remember a similar bewilderment with the choice of the exponent law as the property to be preserved when deriving the definition of fractional or negative exponents.)

Particularly with regard to the compound naming structure that I have been discussing above (though the process may also be at work in some polysyllabic single word names), coining a new term, say *concatenation* for matrix multiplication, results in differing sets of expectations coming to mind. With matrix multiplication, so many of the properties of earlier multiplications do not hold that one wonders about the reason for the choice. To this extent naming is far from psychologically arbitrary and the question of *justification* for the use of a particular compound expression of this sort rather than a new term being coined can be raised.

(The history of conic sections provides an earlier example of shift of attention signalled by a change in the name. The earlier names for the various sections were based on the type of angle of the cones from which the particular curves were obtained; thus *oxytome* (acute-cut), orthotome (right-cut) and *amblytome* (obtuse-cut) were the mathematical names. Apollonius renamed the curves as a result of discovering how they might all be derived from a single cone, using words that had previously been used in the pre-Euclidean dissection theory of area, namely *ellipsis* (the falling-short one), *parabola* and *hyperbola*. Thus the name now reflected the *symptom* of each curve, that is the specifying condition or property of the curve itself, rather than some classificatory naming to do with its construction.)

At some important level, this article has to do with naming. "A rose, by any other name, would still have thorns" suggests that Shakespeare's related observation was a triviality in the sense that the properties of an object are not affected by the choice of name. But on the other hand, there is a sense in which the observation is actually false, and *that* has to do with saliency. Naming can frequently stress certain features and ignore others by means of what it draws attention to. The force of the name "spherical triangle" came from the similarity of the role played by the concept so named in the theory of spherical geometry to that of (planar) triangle in the plane theory.

The power of this metaphoric adjectival construct is that it creates links between the new and the old setting, by highlighting (or, from an alternative philosophical position, creating) certain commonalities. The extended meaning may also have a certain retrospective effect on the old setting, by drawing attention to certain features of it that may not have been noticed previously or considered of importance. It certainly has the effect of altering the balance between the various connotations of the concept.

The use of alternative forms of naming rather than coining new words has in part to do with linguistic conservatism, including a reluctance to create a large number of neologisms. The price paid is that of concept-stretching, a relative commonplace to which mathematical neophytes have to become accustomed. One place in which mathematical literary criticism can help them is by providing a language with which to talk about such processes, as well as more straightforward examples of the process at work in everyday language, which may not be so conceptually demanding.

Note

Colette Laborde has pointed out to me that two terms central to the discipline of *didactique de mathematiques*, namely *contrat didactique* and *transposition didactique* (which have the same structural form as the examples I have discussed in this paper), have had a difficult reception, in part precisely because of crossed interpretations. At one level, the last thing that a *contrat didactique* is is a contract (because it is completely tacit and unspoken); a *transposition didactique* usually involves a far greater alteration of material, structure and form than does a musical transposition.

Bibliography

Fauvel, J. [1988] Cartesian and Euclidean rhetoric *For the Learning of Mathematics* 8(2). pp 25-29

Fowler, D. H. [1973] *Introducing real analysis*. Transworld London

Fraleigh, J. B. [1967] *A first course in abstract algebra*. Addison-Wesley Reading, Massachusetts

Halliday, M. A. K. [1975] Some apsects of sociolinguistics, in *Interactions between linguistics and mathematical education*. UNESCO, Copenhagen, pp 64-73

Lakatos, I. [1976] *Proofs and refutations*. Cambridge University Press Cambridge

Lakoff, G. and Johnson, M [1980] *Metaphors we live by* University of Chicago Press, Chicago

Leron, U [1983] Structuring mathematical proofs. *American Mathematical Monthly* 90(3) pp 174-184

Leron, U. [1985] Heuristic presentations: the role of structuring. *For the Learning of Mathematics*, 5(3), pp 7-13

Nolder, R [1985] Complex numbers, *Mathematics Teaching* 110 March, p. 34

Pimm, D [1984] Who is we? *Mathematics Teaching* 107. June, pp 39-42

Pimm. D. [1987] *Speaking mathematically: communication in mathematics classrooms*. Routledge & Kegan Paul, London

Russell, B [1970, first edn 1912] *The problems of philosophy* Oxford University Press, Oxford

Sykes, J. B. (ed.) [1982] *The concise Oxford dictionary* Oxford University Press, Oxford

Winner, E. [1979] New names for old things: the emergence of metaphoric language *Journal of Child Language* 6 pp 469-491

The essence of the type [of liberal education] is a large discursive knowledge of the best literature. The ideal product of the type is the man who is acquainted with the best that has been written. He will have acquired the chief languages, he will have considered the histories of the rise and fall of nations, the poetic expression of human feeling, and have read the great dramas and novels. He will also be well grounded in the chief philosophies, and have attentively read those philosophic authors who are distinguished for lucidity of style

It is obvious that, except at the end of a long life, he will not have much time for anything else if any approximation is to be made to the fulfilment of this programme. One is reminded of the calculation in a dialogue of Lucian that, before a man could be justified in practising any of the current ethical systems, he should have spent a hundred and fifty years in examining their credentials.

A. N. Whitehead
